EDITORIAL

Genetic commonality between inflammatory bowel disease and sarcoidosis: the beginning of the end or the end of the beginning?

S. Cerri*, R.M. du Bois# and P. Spagnolo*

Crohn's disease (CD) and ulcerative colitis, collectively known as inflammatory bowel disease (IBD), and sarcoidosis are multifactorial disorders thought to result from complex interactions between environmental stimuli (*e.g.* infectious agents), susceptibility genes (which may predispose to the development of granulomatous inflammation) and modifier genes (which may affect disease phenotype in people already susceptible). Neither IBD nor sarcoidosis is the result of defects in a single major gene or chemical pathway; instead, multiple genes, each contributing a relatively minor effect, are likely to be involved. In addition to the granulomatous histopathology, both diseases share a number of similarities in terms of ocular, dermatological and joint manifestations, although sarcoidosis rarely involves the gastrointestinal tract and IBD rarely involves the lung. Immunological, bacteriological and genetic data support a link between CD and sarcoidosis. Both disorders share a similar, yet distinct, immune response, histologically defined by non-caseating granulomas. Up to 50% of patients with CD have been reported to test positive for Kveim antigens, although these data are not replicated in all studies [1]. Mycobacteria have been detected in intestinal tissue of patients with CD and in the lungs of patients with sarcoidosis [2, 3]. A CD4/CD8 lymphocyte ratio >3.5 on bronchoalveolar lavage (BAL) is commonly observed in sarcoidosis, often preceding granuloma formation [4], with a similar expansion of T-cell subsets demonstrated in the lungs of patients with CD [5]. CD4 lymphocytes are activated in the intestinal mucosa of patients with CD, suggesting a cell-mediated disease mechanism in the gut similar to that displayed in pulmonary sarcoidosis. Furthermore, both in sarcoidosis and CD the concept of a genetic component is supported by disease being more common in monozygotic twins compared with dizygotic twins. In contrast with these similarities, tuberculin anergy, salivary gland involvement, the classical lymph and salivary gland patterns of 67-gallium uptake and increased serum angiotensin-converting enzyme, observed in sarcoidosis but not in IBD, are discriminating factors.

In the current issue of the *European Respiratory Journal*, based on the hypothesis that sarcoidosis and IBD may share susceptibility loci, FISCHER *et al.* [6] tested 106 single nucleotide polymorphisms (SNPs) known or suggested to be associated with IBD for a potential association with sarcoidosis and its phenotypes. The study population included a large cohort of sarcoidosis patients and matched controls together with 342 trios (affected offspring and both parents), all from Germany. The non-synonymous SNP rs11209026 (arginine to glutamine substitution at residue 381 (Arg381Gln)) within the interleukin (IL)23 receptor (*IL23R*) gene was associated with chronic sarcoidosis, with similar results observed in a transmission disequilibrium test analysis in an independent family sample. Two separate markers, rs12035082 located at chromosome 1q24.3 and rs916977 located within hect domain and RLD 2 (*HERC2*) gene on chromosome 15q13.1, were associated with acute sarcoidosis and with sarcoidosis as a whole, respectively.

This article raises a number of issues that are important for understanding how genetic data from patients with complex diseases are to be interpreted and how genetic findings might be applied to gain a better understanding of those (often common) factors that drive the pathogenetic process. These issues include: an appreciation of the need for validation studies in support of novel genetic findings; how to interpret the data with regard to the identification of the location of the causal variant(s); how to assign functional significance to genetic observations; and the implications of the findings for other granulomatous diseases and clinical practice.

While the findings by FISCHER *et al.* [6] look robust and intriguing, being derived from SNP analyses from an impressive number of individuals, there are always concerns about the reproducibility of the data in association studies when there is no validation cohort, particularly when other loci of interest have been identified by this and other groups. Therefore, a genetic association, regardless of its strength and biological plausibility, should always be considered the end of the beginning. How certain is it that reported associations between polymorphisms and disease are real? Confidence is increased by the findings being replicated in independent patient populations, ideally from different ethnicities. In fact, population-specific gene–gene or gene–environment interactions

*Center for Rare Lung Diseases, Dept of Oncology, Haematology and Respiratory Diseases, University of Modena and Reggio Emilia, Modena, Italy. #Imperial College, London, UK. ¶Authors are listed alphabetically.

CORRESPONDENCE: P. Spagnolo, Center for Rare Lung Diseases, Dept of Oncology, Haematology and Respiratory Diseases, University of Modena and Reggio Emilia, Via del Pozzo 71, Modena 41100, Italy. E-mail: paolo.spagnolo@unimore.it

may account for the failure to convincingly confirm or refute the majority of associations reported to date [7]. In this regard the findings of FISCHER *et al.* [6] will assume greater importance when their findings are confirmed by others.

How certain can we be that the SNPs that have been reported to be associated with the acute and chronic sarcoidosis phenotypes are causative rather than in linkage with the true causative locus? The mere presence of a predicted deleterious or protective non-synonymous-associated polymorphism does not discount the possibility that a nearby SNP in linkage disequilibrium (defined as the tendency for genetic variants located in close proximity on the same chromosome to occur together more often than expected by chance) may be the true causative variant, may confer an independent molecular mechanism for the phenotype, or may interact epistatically with the non-synonymous variant to cause disease (fig. 1). Sequencing the region of interest should allow causal allele(s) to be identified.

The functional relevance of the associations reported by FISCHER *et al.* [6] can only be speculated upon given the data presented in this manuscript. IL23 stimulates survival and proliferation of IL17-producing CD4+ T-helper (Th17) cells, thus serving as a key master cytokine regulator. The Gln allele at residue 381 in the IL23R protein confers significant protection against CD [8]. Conversely, the more common allele at position 381 is an Arg residue (occurring at a frequency of ~95%) highly conserved across multiple species. While the mechanistic impact of the *IL23R* 381 Gln residue remains unclear, interestingly, this genetic variant also exerts a similar protective effect against psoriasis, which is consistent with the clinical observation that psoriasis and CD frequently occur together [9].

The IL23/Th17 pathway has recently been suggested to be involved in a number of chronic inflammatory diseases other than IBD [10–12]. While to date there are no studies
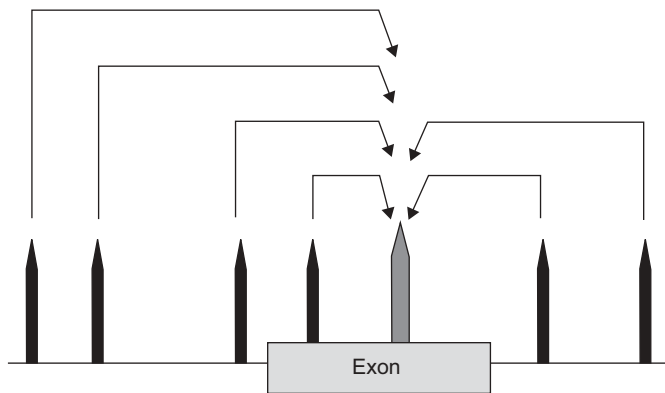
FIGURE 1. Linkage disequilibrium (LD) is the nonrandom segregation or co-inheritance of genetic markers in a population. LD is often the result of close physical proximity. A single-nucleotide polymorphism (SNP) at a certain locus may be associated with a specific phenotype (*i.e.* acute or chronic sarcoidosis) either because it is causative itself or because it is in LD with the disease causing SNP. In this scheme the true/biologically causal variant is the one shown as ▮, while the SNPs shown as ▮ are associated not because they are also biologically causal but indirectly (*i.e.* because of the LD with the causal variant).

investigating the role of Th17 cells in sarcoidosis pathogenesis, it is conceivable that an imbalance between effector T-cells, such as Th17 cells, and CD4+ T-regulatory subsets could lead to or contribute to an exuberant immune response and granulomatous inflammation characteristic of sarcoidosis. In the study by FISCHER *et al.* [6] IL23R was not differentially expressed in BAL cells from sarcoidosis patients as compared with healthy individuals, although this finding does not rule out a causative role for IL23R as the associated genetic variant(s) may influence gene function rather than mRNA expression. Less biologically plausible are the associations between markers rs916977 and sarcoidosis as a whole and between rs12035082 and acute sarcoidosis, the simplest explanation being that these variants may be in linkage disequilibrium with another functional polymorphism located nearby. At present, the function of the associated polymorphisms and the mechanisms by which they contribute to disease pathogenesis are unknown, thus hampering the interpretation of the results.

Despite these concerns about reproducibility and functional relevance of their findings, the observations by FISCHER *et al.* [6] are intriguing by suggesting that there may be a commonality to the immune and/or inflammatory processes that result in granuloma formation, irrespective of the exact disease that is manifest, which may be regulated genetically. Granulomas have a protective role in many diseases caused by an infective agent by preventing the spread of pathogens, including bacteria, mycobacteria, fungi, viruses, protozoans and helminths. In the lung, granulomas are essential in controlling mycobacterial infections, tuberculosis being the commonest, although leprosy, leishmaniasis and schistosomiasis also elicit pulmonary and nonpulmonary granulomatous reactions. However, the role of the granuloma in diseases of unknown aetiology is unclear. For instance, in pulmonary diseases such as sarcoidosis and Wegener's granulomatosis, and in nonpulmonary diseases such as CD, the granulomatous inflammation accounts for much of the pathology, and the protective benefits, if any, are not known. The similarities in the histopathological and immunological features of granulomatous disorders suggest that, whatever the specific aetiology, they may share, at least in part, a common genetic background. Nonetheless, there is also evidence that this may not be necessarily the case. The chronic beryllium disease (CBD) model provides the best example of a known environmental trigger (beryllium) interacting with a well-defined genetic predisposition (the HLA-DPB1 Glu[69] allele) leading to a well-characterised disease state [13]. Sarcoidosis may be clinically indistinguishable from CBD but is not associated with the Glu[69] allele [14, 15]. However, *NOD2* (nucleotide-binding oligomerisation domain protein 2), which encodes an intracellular protein involved in innate immunity through recognition of bacterial pathogen-associated molecular patterns, is a major susceptibility gene for CD [16, 17] but not for sarcoidosis [18, 19], although patients carrying the functional 2104 (702W) polymorphism are more likely to develop severe pulmonary disease, as defined by radiographic stage and pulmonary function data [20]. Therefore, lessons learnt about the genetic influences on one granulomatous disorder may provide insights into the genetics and immunopathology of other similar ones and the findings of FISCHER *et al.* [6] that there are polymorphisms common to CD and sarcoidosis would support this view. Nonetheless, the relevance of the associations

reported by FISCHER *et al.* [6] to granulomatous diseases other than sarcoidosis needs further study.

What is the potential clinical relevance of the findings of FISCHER *et al.* [6]? Much of the interest surrounding genetic association studies centres on the potential clinical application of polymorphisms that serve as markers for disease. However, the low sensitivity and specificity of any individual marker, the modest risk conferred by each individual allele as well as the high prevalence of the risk alleles in the normal population argue against the use of these genetic data either in disease diagnosis or in patient classification. In contrast, the elucidation of the genetic background of granulomatous diseases has great potential to expand our understanding of disease immunopathogenesis. With that greater understanding may come novel and better treatments and more tailored therapy.

The genetic basis of common diseases remains largely unknown. While technological advances in genotyping and the potential of genome-wide association analysis have led to a rapid increase in the number of suggested susceptibility genes, most of the studies lack crucial information on the mechanisms through which the polymorphisms could affect an individual's susceptibility to disease. The study by FISCHER *et al.* [6] suggests the existence of a common genetic basis, and possibly of overlapping pathways, in IBD and sarcoidosis, and supports the notion that genetic factors may not only confer susceptibility to sarcoidosis but also contribute to its wide variety of clinical manifestations and prognoses, which range from acute uncomplicated to chronic and progressive multisystem disease [21]. Emphasis must now be placed on showing that these genetic associations are reproducible in sarcoidosis cohorts other than the German patients studied in this article and in understanding their functional consequences.

STATEMENT OF INTEREST
None declared.

REFERENCES
1 Mitchell DN, Cannon P, Dyer NH, *et al.* Further observations on Kveim test in Crohn's disease. *Lancet* 1970; 2: 496–498.
2 Xavier RJ, Podolsky DK. Unravelling the pathogenesis of inflammatory bowel disease. *Nature* 2007; 448: 427–434.
3 Song Z, Marzilli L, Greenlee BM, *et al.* Mycobacterial catalase-peroxidase is a tissue antigen and target of the adaptive immune response in systemic sarcoidosis. *J Exp Med* 2005; 201: 755–767.
4 Drent M, Mansour K, Linssen C. Bronchoalveolar lavage in sarcoidosis. *Semin Respir Crit Care Med* 2007; 28: 486–495.
5 Smiéjan JM, Cosnes J, Chollet-Martin S, *et al.* Sarcoid-like lymphocytosis of the lower respiratory tract in patients with active Crohn's disease. *Ann Intern Med* 1986; 104: 17–21.
6 Fischer A, Nothnagel M, Franke A, *et al.* Association of inflammatory bowel disease risk loci with sarcoidosis and its acute and chronic subphenotypes. *Eur Respir J* 2011; 37: 610–616.
7 Hirschhorn JN, Lohmueller K, Byrne E, *et al.* A comprehensive review of genetic association studies. *Genet Med* 2002; 4: 45–61.
8 Duerr RH, Taylor KD, Brant SR, *et al.* A genome-wide association study identifies IL23R as an inflammatory bowel disease gene. *Science* 2006; 314: 1461–1463.
9 Cargill M, Schrodi SJ, Chang M, *et al.* A large-scale genetic association study confirms IL12B and leads to the identification of IL23R as psoriasis-risk genes. *Am J Hum Genet* 2007; 80: 273–290.
10 Guttman-Yassky E, Lowes MA, Fuentes-Duculan J, *et al.* Low expression of the IL-23/Th17 pathway in atopic dermatitis compared to psoriasis. *J Immunol* 2008; 181: 7420–7427.
11 Abraham C, Cho JH. IL-23 and autoimmunity: new insights into the pathogenesis of inflammatory bowel disease. *Annu Rev Med* 2009; 60: 97–110.
12 Paust HJ, Turner JE, Steinmetz OM, *et al.* The IL-23/Th17 axis contributes to renal injury in experimental glomerulonephritis. *J Am Soc Nephrol* 2009; 20: 969–979.
13 Richeldi L, Sorrentino R, Saltini C. HLA-DPB1 glutamate 69: a marker of beryllium disease. *Science* 1993; 262: 242–244.
14 Foley PJ, Lympany PA, Puscinska E, *et al.* Analysis of MHC encoded antigen-processing genes TAP1 and TAP2 polymorphisms in sarcoidosis. *Am J Respir Crit Care Med* 1999; 160: 1009–1014.
15 Maliarik MJ, Chen KM, Major ML, *et al.* Analysis of HLA-DPB1 polymorphisms in African-Americans with sarcoidosis. *Am J Respir Crit Care Med* 1998; 158: 111–114.
16 Ogura Y, Bonen DK, Inohara N, *et al.* A frameshift mutation in NOD2 associated with susceptibility to Crohn's disease. *Nature* 2001; 411: 603–606.
17 Hugot JP, Chamaillard M, Zouali H, *et al.* Association of NOD2 leucine-rich repeat variants with susceptibility to Crohn's disease. *Nature* 2001; 411: 599–603.
18 Milman N, Nielsen OH, Hviid TV, *et al.* CARD15 single nucleotide polymorphisms 8, 12 and 13 are not increased in ethnic Danes with sarcoidosis. *Respiration* 2007; 74: 76–79.
19 Schürmann M, Valentonyte R, Hampe J, *et al.* CARD15 gene mutations in sarcoidosis. *Eur Respir J* 2003; 22: 748–754.
20 Sato H, Williams HRT, Spagnolo P, *et al.* CARD15/NOD2 polymorphisms are associated with severe pulmonary sarcoidosis. *Eur Respir J* 2010; 35: 324–330.
21 Baughman RP, Lower EE, du Bois RM. Sarcoidosis. *Lancet* 2003; 361: 1111–1118.